数理解析研究所講究録
72巻 1969年 1-34

代数方程式の数値解法

日本建設コンサルタント　平野菅保

数値計算は有限桁で行なわれる。この点が非常に重要な点であり、その結果従来の常識では考えられない様な解を得る事がある。

まずその第1は、与えられた代数方程式の係数の精度(桁数)に比較して、得られる解の精度(桁数)が悪い事がある点である。これは数値計算の計算方法の善し悪しには関係のない事で、誤差を含む代数方程式の係数が与えられた時、すでに解に入ってしまう誤差である。例えば、与えられた代数方程式の係数の精度(桁数)がすべてm桁あり、その代数方程式がP重根を含んでいれば、そのP重根の解はm/P桁の精度しか得ることが出来ない。又近接根を含む代数方程式の場合も重根と同様な傾向がある。但し、代数方程式の係数の誤差間に何らかの関係がある場合には、解の精度がm/P桁よりよくなる事があり、この事が桁数を多くして計算をする

理由の１つでもある．

第２は，１つの解が得られて，その解を除いた次数の１次下がった代数方程式の係数を求める時，計算方法が悪いと求められた解の誤差により１次下がった代数方程式の係数に誤差が入ってしまう．これは各解によって係数に要求する精度（桁数）が係数毎に異なるからである．もし一度でも，次数を下げた代数方程式の係数に誤差が入ると，その代数方程式によって得られる解は，初めに与えられた代数方程式の解とは異なった解，即ち本来得られる筈の解の精度（桁数）が得られなくなってしまう．しかしこれも又桁数を多くして計算を行なえば防ぐ事が出来る．

次に，解へ段々と収束させていく解法，ニュートン法，ベアストー法（ヒッチコック法）では，補正値を求めるために１次微分しか用いていない点であるが，これは補正値の絶対値が充分に小さいと仮定した場合の事であって，実際の計算での補正値の絶対値はその様に充分小さいとは限らない．従って２次微分以上の項を考慮して補正値を求めないと，いわゆるニュートン法に於ける飛ぶとゆう現象が起こる．この現象はせっかく解の近くまで収束していながら，補正値を加えると，かえって解より離れた近似解になってしまい，時によっては異なった値の他の解に近づいてしまう事もある．

(1)　代数方程式の係数の誤差と解の誤差

係数に誤差を含まない時でも，有限固定桁で計算を行なうと，必ず数値に誤差が入つてしまう．これは計算に用いる桁数より下位の桁は零でなく任意の数値でも同じ数値になる．即ち誤差である．従つて与えられた代数方程式の係数には必ず誤差が含まれている．

$$f(Z)=\sum_{i=0}^{n}(C_i \pm \Delta C_i)Z^i=0 \qquad (1-1)$$

n 次の代数方程式を（1-1）式で表わし，近似解 $\bar{Z}$ を（1-1）式に代入し，$f(\bar{Z})$ が誤差のみになつたならば，近似解 $\bar{Z}$ をそれ以上真の解へ近づける事は出来なくなる．例えば

$$max_{i=0,1,2,\cdots,n}(|\Delta C_i \bar{Z}^i|) \geq |f(\bar{Z})| \qquad (1-2)$$

（1-2）式の條件を満足したならば，更に計算を行なつても解の精度は良くならず，意味がない．しかし

$$(Z-Z_1)^n=0 \qquad (1-3)$$

（1-3）式の Z_1 に誤差を入れて

$$(Z-(Z_1+\Delta Z_1))^n=0 \qquad (1-4)$$

（1-4）式で桁数を多くとつて計算し，（1-1）式の形にすると，

$$C_n \pm \Delta C_n = 1.0$$

$$C_{n-1} \pm \Delta C_{n-1} = -n(Z_1+\Delta Z_1)$$

$$\vdots$$

$$C_1 \pm \Delta C_1 = (-1)^{n-1} n (Z_1 + \Delta Z_1)^{n-1}$$

$$C_0 \pm \Delta C_0 = (-1)^{n} (Z_1 + \Delta Z_1)^{n} \qquad (1-5)$$

（1-5）式の係数の誤差間は独立ではなく、（1-2）式よりも更に解の精度をあげる事が出来る。

次に、（1-1）式の解Z_1がP重根であるとする。（1-1）式を点Z_1でテーラー展開してみると

$$f(Z) = \sum_{i=p}^{n} \{ f^{(i)}(Z_1)/i! \}(Z - Z_1)^{i} = 0 \qquad (1-6)$$

［Z_1はP重根であるから　$f(Z_1) = f'(Z_1) = \cdots\cdots = f^{(p-1)}(Z_1) = 0$］

$|Z - Z_1|$が充分小さい値の時、（1-6）式より、$|f(Z)|$は$|Z - Z_1|$に対してP乗に比例している。一方、$|Z - Z_1|$が充分小さい時、ZとZ_1の数値を比較してみると、上位の桁の数値は等しく、下位の桁の数値のみ異なつているから、（1-1）式の各項の絶対値は、$|Z - Z_1|$が 1/10、1/100と小さくなつても、ほとんど変化しない。従つて（1-1）式と（1-6）式にZを代入した時は等しいので、$|Z - Z_1|$が1桁小さくなる様なZを（1-1）式に代入すれば、即ち近似解の精度が1桁良くなれば、$|f(Z)|$はP桁小さくなり、有効桁数はP桁失われる。この事から「与えられた代数方程式の係数がすべてm桁の精度をもつならば、P重根の解の精度は m/P 桁ある。」といえる。

例として．2重根の解を含む3次代数方程式を示す．

$$f(Z) = C_3 Z^3 + C_2 Z^2 + C_1 Z + C_0$$

$C_3 = 1.0000000$　　　$C_1 = 15.000000$

$C_2 = -7.0000000$　　　$C_0 = -9.0000000$

真の解は　$Z_1 = 1.0000000$　　　$Z_2 = Z_3 = 3.0000000$

$Z = 3.0003000$	$Z = 3.0030000$
$C_0 = -9.0000000$	$C_0 = -9.0000000$
$C_1 Z = 45.004500$	$C_1 Z = 45.045000$
$C_2 Z^2 = -63.012601$	$C_2 Z^2 = -63.126063$
$C_3 Z^3 = 27.008101$	$C_3 Z^3 = 27.081081$
$f(Z) = 0.000000\underline{0}$	$f(Z) = 0.000018\underline{0}$

$Z = 3.0300000$	$Z = 3.3000000$
$C_0 = -9.0000000$	$C_0 = -9.0000000$
$C_1 Z = 45.450000$	$C_1 Z = 49.500000$
$C_2 Z^2 = -64.266300$	$C_2 Z^2 = -76.230000$
$C_3 Z^3 = 27.818127$	$C_3 Z^3 = 35.937000$
$f(Z) = 0.001827\underline{0}$	$f(Z) = 0.207000\underline{0}$

関数値$f(Z)$は10^{-7}の桁が誤差＿になっている。又．解3.0は2重根であるから．Zの精度が1桁良くなると．関数値$f(Z)$の有効桁数は2桁減少している事がわかる。従って8桁で計算をすると．解3.0は4桁目まで決定出來る．

次に、係数の誤差間が独立でない場合の例(1−5)式を考えてみる。

$$n = 3 \qquad Z_1 = 3.14159265$$

$$\Delta Z_1 = -0.00100000 \qquad Z_1 + \Delta Z_1 = 3.14059265$$

とすると、$Z_1 + \Delta Z_1$ が Z_1 に対して3桁の精度であるから、3桁で(1−5)式の係数を求め、次に $Z_1 + \Delta Z_1$ を求める計算を行なう。$Z_1 + \Delta Z_1$ は3重根であるから、$Z_1 + \Delta Z_1$ の精度は1桁であり、真の解 Z_1 に対しても1桁の精度しか得られない。

$$C_3 \pm \Delta C_3 = 1.0 = 1.00$$

$$C_2 \pm \Delta C_2 = -3.0\,(Z_1 + \Delta Z_1) = -9.42$$

$$C_1 \pm \Delta C_1 = 3.0\,(Z_1 + \Delta Z_1)^2 = 29.6$$

$$C_0 \pm \Delta C_0 = -1.0\,(Z_1 + \Delta Z_1)^3 = -31.0$$

Z	2.60	3.10
$(C_3 \pm \Delta C_3)\,Z^3$	17.6	29.8
$(C_2 \pm \Delta C_2)\,Z^2$	−63.7	−90.5
$(C_1 \pm \Delta C_1)\,Z$	77.0	91.8
$(C_0 \pm \Delta C_0)$	−31.0	−31.0
$f(Z)$	0.1	0.1

Z が2.6あるいは3.1でも $f(Z)$ は殆んど誤差のみになる。

ところが(1−5)式の係数を求めるのに次の様に9桁で計算し、次に $Z_1 + \Delta Z_1$ を求めると、$Z_1 + \Delta Z_1$ の精度は3桁であり

真の解Z_1に対しても3桁の精度が得られるのである。

$$C_3 \pm \Delta C_3 = \quad 1.0 \qquad\qquad = \quad 1.00000000$$

$$C_2 \pm \Delta C_2 = -3.0\,(Z_1 + \Delta Z_1) = -\ 9.42177795$$

$$C_1 \pm \Delta C_1 = \quad 3.0\,(Z_1 + \Delta Z_1)^2 = \quad 29.5899666$$

$$C_0 \pm \Delta C_0 = -1.0\,(Z_1 + \Delta Z_1)^3 = -30.9766772$$

Z	3.13500000	3.14400000
$(C_3 \pm \Delta C_3)Z^3$	30.8114854	31.0776100
$(C_2 \pm \Delta C_2)Z^2$	−92.5993536	−93.1317877
$(C_1 \pm \Delta C_1)Z$	92.7645453	93.0308550
$(C_0 \pm \Delta C_0)$	−30.9766772	−30.9766772
$f(Z)$	− 0.0000001	0.0000001

計算して得られる解の誤差は、どうしたら得られるか。真の解Z_{Tj}を(1−1)式に代入してみる。

$$f(Z_{Tj}) + \varepsilon_j = \sum_{i=0}^{n} (C_i \pm \Delta C_i) Z_{Tj}^i + \varepsilon_j = 0$$

$$j = 1, 2, \cdots\cdots, n \qquad (1\text{−}8)$$

ε_jなる誤差項を加えないと、等号は成り立たない。ε_jとしては

$$\varepsilon_j = \max_{i=0,1,2,\cdots\cdots,n} \left\{ \left| (\pm \Delta C_i) Z_{Tj}^i \right| \right\}$$

$$= \left| (\pm \Delta C_m) Z_{Tj}^m \right| \qquad j = 1, 2, \cdots\cdots, n \qquad (1\text{−}9)$$

$$\varepsilon_j = \sum_{i=0}^{n} \left| (\pm \Delta C_i) Z_{Tj}^i \right| \qquad j = 1, 2, \cdots\cdots n \qquad (1\text{−}10)$$

(1−10)式でε_jを求めると、これ以下で述べる解の誤差を求める方法では、解の誤差が大きめになるため、(1−9)式を

採用する．又（1-1）式を解くと

$$f(Z_{Tj}\pm\Delta Z_{Tj})=\sum_{i=0}^{n}(C_i\pm\Delta C_i)(Z_{Tj}\pm\Delta Z_{Tj})^i=0$$

$$j=1,2,\cdots\cdots,n \qquad (1-11)$$

$Z_{Tj}\pm\Delta Z_{Tj}$ が得られる．次に

$$y=Z/Z_{Tj}\quad,\quad \Delta y_{Tj}=\Delta Z_{Tj}/Z_{Tj}$$

$$j=1,2,\cdots\cdots,n \qquad (1-12)$$

と変数変換をして，（1-8），（1-11）式をそれぞれ

$$f(Z_{Tj}\cdot y)+\varepsilon_j=\sum_{i=0}^{n}(C_i\pm\Delta C_i)(Z_{Tj}\cdot y)^i+\varepsilon_j=0$$

$$j=1,2,\cdots\cdots,n \qquad (1-13)$$

$$f(Z_{Tj}\cdot(y\pm\Delta y_{Tj}))=\sum_{i=0}^{n}(C_i\pm\Delta C_i)(Z_{Tj}\cdot(y\pm\Delta y_{Tj}))^i=0$$

$$j=1,2,\cdots\cdots,n \qquad (1-14)$$

で表わし，$y=1.0$ であるから，（1-14）式から（1-13）式を減ずると

$$\sum_{i=1}^{n}(\overset{(i)}{f}(Z_{Tj})/i!)(\pm\Delta y_{Tj})^i=\varepsilon_j$$

$$j=1,2,\cdots\cdots,n \qquad (1-15)$$

（1-15）式の左辺の各項は，どの項も誤差であるから

$$|(\overset{(i)}{f}(Z_{Tj})/i!)(\pm\Delta y_{Tj})^i|\le\varepsilon_j$$

$$i=1,2,\cdots\cdots,n\ ;\quad j=1,2,\cdots\cdots,n \qquad (1-16)$$

であると考える．（1-16）式を満足する $|\pm\Delta y_{Tj}|$ を求めるために

$$|\pm\Delta y_{Tj}|=min_{i=1,2,\cdots\cdots,n}\left(\left|\sqrt[i]{\frac{\varepsilon_j\cdot i!}{\overset{(i)}{f}(Z_{Tj})}}\right|\right)$$

$$= \left| \sqrt[\overline{m}]{\frac{\varepsilon_j \cdot \overline{m}!}{f^{(\overline{m})}(Z_{Tj})}} \right| \qquad j=1,2,\cdots,n \qquad (1\text{-}17)$$

(1-17) 式で求めた $|\pm\Delta y_{Tj}|$ を (1-16) 式の左辺に代入する．まず $i=\overline{m}$ の項は

$$\left| \frac{f^{(\overline{m})}(Z_{Tj})}{\overline{m}!} \left(\sqrt[\overline{m}]{\frac{\varepsilon_j \cdot \overline{m}!}{f^{(\overline{m})}(Z_{Tj})}} \right)^{\overline{m}} \right| = \varepsilon_j$$

$$j=1,2,\cdots,n \qquad (1\text{-}18)$$

他の項は

$$\left| \frac{f^{(i)}(Z_{Tj})}{i!} \left(\sqrt[\overline{m}]{\frac{\varepsilon_j \cdot \overline{m}!}{f^{(\overline{m})}(Z_{Tj})}} \right)^{i} \right| \qquad j=1,2,\cdots,n$$

$$\leq \left| \frac{f^{(i)}(Z_{Tj})}{i!} \left(\sqrt[i]{\frac{\varepsilon_j \cdot i!}{f^{(i)}(Z_{Tj})}} \right)^{i} \right| = \varepsilon_j \qquad (1\text{-}19)$$

即ち，(1-17) 式で求めた $|\pm\Delta y_{Tj}|$ は必ず (1-16) 式を満足している．従って (1-17) 式から解の相対誤差を求める事が出来る．解の絶対誤差は (1-12) 式を用い，Z_{Tj} の代りに $Z_{Tj}\pm\Delta Z_{Tj}$ を代入し

$$|\pm\Delta Z_{Tj}| = |Z_{Tj}\pm\Delta Z_{Tj}| \min_{i=1,2,\cdots,n} \left(\left| \sqrt[i]{\frac{\varepsilon_j \cdot i!}{f^{(i)}(Z_{Tj}\pm\Delta Z_{Tj})}} \right| \right)$$

$$j=1,2,\cdots,n \qquad (1\text{-}20)$$

で求められる．又重根である解，あるいは近接根を含む解を求める時は (1-20) 式で $i=1$ が選ばれない．例えば，3重

根の解の場合は $i=3$ が選ばれる．次に，3重根の解を含む3次代数方程式を例として，（1−20）式で計算した誤差をあげる．

$$f(z)=(\pi+ei)\{z-(\pi+ei)\}^3=C_3z^3+C_2z^2+C_1z+C_0$$

上式の C_3，C_2，C_1，C_0 を計算し，誤差を入れる。

$$C_3\pm\Delta C_3=\quad 3.14159210+\quad 2.71828130\,i$$

$$C_2\pm\Delta C_2=-\ 7.44164200-\ 51.2384030\,i$$

$$C_1\pm\Delta C_1=-115.901801\ +181.198677\,i$$

$$C_0\pm\Delta C_0=\ 285.555090\ -\ 84.7328600\,i$$

$\pm\Delta C_3=5.5\times10^{-7}+5.2\times10^{-7}\,i$ 　　 $|\pm\Delta C_3|=7.5690\times10^{-7}$

$\pm\Delta C_2=2.9\times10^{-6}+2.3\times10^{-6}\,i$ 　　 $|\pm\Delta C_2|=3.7014\times10^{-6}$

$\pm\Delta C_1=8.1\times10^{-6}+8.8\times10^{-6}\,i$ 　　 $|\pm\Delta C_1|=1.1960\times10^{-5}$

$\pm\Delta C_0=3.2\times10^{-5}+3.2\times10^{-5}\,i$ 　　 $|\pm\Delta C_0|=4.5255\times10^{-5}$

$$z_1=3.15217390+2.74294324\,i$$

$$z_2=3.15749449+2.69689966\,i$$

$$z_3=3.11511032+2.71500419\,i$$

$$|\Delta z_1|=0.0072310$$

$$|\Delta z_2|=0.0071916$$

$$|\Delta z_3|=0.0071117$$

真の誤差より（1−20）式で計算した誤差は小さめにでている．これは ε_j を（1−9）式により計算しているからであり，ε_j

を（1−10）式より計算すると誤差は大きめにでる。

（2）　解への収束方法

与えられた代数方程式を

$$f(Z) = C_0 + \sum_{i=1}^{n} C_i Z^i = 0 \qquad (2-1)$$

とする．繰り返し計算の j 回目の近似値を Z_j とすると

$$f(Z) = C_{0j} + \sum_{i=1}^{n} C_{ij} (Z - Z_j)^i = 0 \qquad (2-2)$$

$$C_{0j} = f(Z_j)$$

$$C_{ij} = f^{(i)}(Z_j) / i!$$

で表わせる。もし（2−2）式で $C_{0j} = 0$ であるならば，Z_j は解である。次に $j+1$ 回目の近似値 $Z_{j+1} = Z_j + \Delta Z_j$ を（2−1）式に代入し

$$|f(Z_{j+1})| = |f(Z_j + \Delta Z_j)| < |f(Z_j)| \qquad (2-3)$$

になる様な ΔZ_j を求め，これを繰り返して $j \to \infty$ とすれば $f(Z_j) \to 0$ となり解が求められる。

$\alpha \geq 1.0$ なる実数 α を用いて（2−2）式を

$$f(Z) = \left(1 - \frac{1}{\alpha}\right) C_{0j} + \frac{1}{\alpha} C_{0j} + \sum_{i=1}^{n} C_{ij} (Z - Z_j)^i = 0 \qquad (2-4)$$

と変形し，（2−4）式の右辺の第2項と第3項の合計を

$$f(Z) = \frac{1}{\alpha} C_{oj} + \sum_{i=1}^{n} C_{ij} (Z - Z_j)^i \qquad (2-5)$$

とする。次に

$$\left| \sqrt[m]{\frac{-C_{oj}}{\alpha C_{mj}}} \right| = \min_{i=1,2,\cdots,n} \left(\left| \sqrt[i]{\frac{-C_{oj}}{\alpha C_{ij}}} \right| \right) \qquad (2-6)$$

でmを求める。（m項として２項以上ある時は，いずれを選んでもよい。）極座標を用いて

$$-\frac{C_{oj}}{\alpha C_{mj}} = r_{\alpha m} e^{i\theta_{\alpha m}} \qquad (2-7)$$

と表わし，（2-7）式を用いて仮のΔZ_jを

$$\Delta Z_{jm'} = \sqrt[m]{r_{\alpha m}}\; e^{i(\theta_{\alpha m} + 2\pi m')/m} \qquad m' = 1, 2, \cdots, m \qquad (2-8)$$

とする。（m'はm個の中で任意の１つを選べばよい。）

ここで，（2-5）式に$Z = Z_j + \Delta Z_{jm'}$を代入する．

$$\begin{aligned} f(Z_j + \Delta Z_{jm'}) &= \frac{1}{\alpha} C_{oj} + \sum_{i=1}^{n} C_{ij} (Z_j + \Delta Z_{jm'} - Z_j)^i \\ &= \frac{1}{\alpha} C_{oj} + \sum_{i=1}^{n} C_{ij} (\Delta Z_{jm'})^i \end{aligned} \qquad (2-9)$$

（2-7）式と（2-8）式より

$$-C_{oj} / (\alpha \cdot C_{mj}) = (\Delta Z_{jm'})^m$$

$$\frac{1}{\alpha} C_{oj} + C_{mj} (\Delta Z_{jm'})^m = 0 \qquad (2-10)$$

従って（2-9）式は

$$f(Z_j + \Delta Z_{jm'}) = \sum_{i=1}^{n'} C_{ij} (\Delta Z_{jm'})^i \qquad (2-11)$$

n'：$1, 2, \cdots, n$ の中からmを除いたもの

となる．又（2−6）式で最小値を選んであるので，（2−9）式の $\Delta Z_{jm'}$ の i 次の項に $\Delta Z_{jm'}$ の値を代入すると

$$\left| C_{ij} (\Delta Z_{jm'})^{i} \right| = \left| C_{ij} \left(\sqrt[m]{\frac{-C_{oj}}{\alpha \cdot C_{mj}}} \right)^{i} \right|$$

$$\le \left| C_{ij} \left(\sqrt[i]{\frac{-C_{oj}}{\alpha \cdot C_{ij}}} \right)^{i} \right| = \left| \frac{-C_{oj}}{\alpha} \right| \qquad (2-12)$$

依って（2−9）式の中には，零次の項の絶対値より，絶対値の大きい項はない。もしここで

$$\left| \frac{1}{\alpha} C_{oj} \right| = \left| f(Z_j) \right| > \left| f(Z_j + \Delta Z_{jm'}) \right| \qquad (2-13)$$

ならば $\Delta Z_{jm'}$ を j 回目の補正値 ΔZ_j として

$$Z_{j+1} = Z_j + \Delta Z_j \quad (\Delta Z_j = \Delta Z_{jm'}) \qquad (2-14)$$

を $j+1$ 回目の近似値とする．

$$\left| \frac{1}{\alpha} C_{oj} \right| = \left| f(Z_j) \right| \le \left| f(Z_j + \Delta Z_{jm'}) \right| \qquad (2-15)$$

ならば α を更に大きくして（2−4）式より繰り返す。（2−6）式で α を大きくすると

$$\left| \sqrt[i]{\frac{-C_{oj}}{\alpha \cdot C_{ij}}} \right| = \left| \sqrt[i]{\frac{1}{\alpha}} \right| \left| \sqrt[i]{\frac{-C_{oj}}{C_{ij}}} \right| \qquad (2-16)$$

i の小さい項は $\left| \sqrt[i]{1/\alpha} \right|$ の小さくなるなり方が，i の大きい項よりも速いので，段々低次の項が選ばれ，極限では零でない一番 i の小さい項が選ばれる。一方，$\Delta Z_{jm'}$ の絶対値も小さくなるので，（2−9）式で考えればわかる様に，（2−13

）式を満足するΔZ_{jm}が存在する．（2−13）式が満足されれば$|f(Z_{j+1})| = |f(Z_j+\Delta Z_j)| = |(1-\frac{1}{\alpha})C_{oj} + \bar{f}(Z_j+\Delta Z_j)|$

$$< |(1-\frac{1}{\alpha})C_{oj} + \frac{1}{\alpha}C_{oj}| = |C_{oj}|$$

$$= |f(Z_j)| \qquad (2-17)$$

従って（2−3）式が満足されるから，$j\to\infty$ で $C_{oj}\to 0$ となり，解が1つ求められる．

次に，この収束方法を用いた例をあげる．

$$Z^6 + 26.417062Z^5 + 162.32544Z^4$$
$$+ 681.48974Z^3 + 477.60822Z^2$$
$$+ 1088.9984Z + 74.312797 = 0$$

解は，$Z_1 = -\ 0.070187322$

$Z_2 = -\ 0.13285772 - 1.3416340\,i$

$Z_3 = -\ 0.13285772 + 1.3416340\,i$

$Z_4 = -\ 3.0725744 - 4.4472419\,i$

$Z_5 = -\ 3.0725744 + 4.4472419\,i$

$Z_6 = -\ 19.936010$

初期値を 0.0 とすると，一般のニュートン法と同様に先ず解Z_1が求められる．解Z_1を除いた5次の代数方程式の係数を求め，初期値を 0.0 とし，次の解Z_2の収束計算を繰り返す．

繰り返し回数

1 $\qquad -1.2568739\,i$

2	$-0.13701416-1.3309150\,i$
3	$-0.13285435-1.3416929\,i$
4	$-0.13285772-1.3416340\,i$

４回で収束する．５次の代数方程式の係数は解Z_1が実数であるからすべて実数である．故に，前記の様に初期値 0.0 では一般のニュートン法を用いると複素解Z_2は求められない．

（３） 多項式の除算

与えられた代数方程式を

$$f(Z)=\sum_{i=0}^{n}C_iZ^i=0 \qquad (3-1)$$

とし，得られた解をZ_1とすると，代数方程式の次数を１次下げる計算をしなければならない．

$$\bar{f}(Z)=\left(\sum_{i=0}^{n}C_iZ^i\right)\Big/\left(Z-Z_1\right)=\sum_{i=0}^{n-1}\bar{C}_iZ^i=0 \qquad (3-2)$$

ところが，得られた解Z_1が与えられた代数方程式の係数に要求する精度（桁数）は，値の異なる他の解がその係数に要求する精度（桁数）とは異なる．今，２つの解を$|Z_1|<|Z_2|$として，（1-2）式により，解Z_1，Z_2の精度に最も影響する項を求める．

$$|\pm\Delta C_{i_1}Z_1^{i_1}|=max_{i=0,1,2,\cdots\cdots,n}\left(|\pm\Delta C_iZ_1^i|\right)$$

$$\left|\pm\Delta C_{i2} Z_2^{i2}\right| = max_{i=0,1,2,\cdots\cdots,n}\left(\left|\pm\Delta C_i Z_2^{i}\right|\right) \quad (3-3)$$

（3−3）式中の $\pm\Delta C_i$ は（3−1）式中の係数 C_i の含む誤差である．（3−3）式より

$$\left|\pm\Delta C_{i1} Z_1^{i1}\right| \geq \left|\pm\Delta C_{i2} Z_1^{i2}\right|$$
$$\left|\pm\Delta C_{i2} Z_2^{i2}\right| \geq \left|\pm\Delta C_{i1} Z_2^{i1}\right| \quad (3-4)$$

（3−4）式より

$$\left|Z_1^{i2-i1}\right| \leq \left|\pm\Delta C_{i1} / \pm\Delta C_{i2}\right| \leq \left|Z_2^{i2-i1}\right| \quad (3-5)$$

$$1.0 \leq \left|Z_2 / Z_1\right|^{i2-i1} \quad (3-6)$$

一方．仮定より $1.0 < |Z_2 / Z_1|$ であるから

$$i2 - i1 \geq 0 \quad \text{即ち} \quad i2 \geq i1 \quad (3-7)$$

従って．絶対値の大きい解は．絶対値の小さい解よりも．与えられた代数方程式の次数の高い項の係数の誤差により．誤差が決定される。次に．$\pi + ei$ の2重根及び $0.1(\pi + ei)$ の2重根を含む4次の代数方程式の例をあげる．

$$f(Z) = C_4 Z^4 + C_3 Z^3 + C_2 Z^2 + C_1 Z + C_0$$

まず．次数の低い項の係数に誤差が多く入っている場合は

$$C_4 = 3.141592\underline{1} + 2.718281\underline{3}\ i$$
$$C_3 = -5.45720\underline{40} - 37.57482\underline{8}\ i$$
$$C_2 = -54.47384\underline{2} + 85.16337\underline{4}\ i$$
$$C_1 = 62.8220\underline{90} - 18.6412\underline{00}\ i$$
$$C_0 = -11.274\underline{100} - 5.100\underline{1000}\ i$$

$$Z_1 = 3.14\underline{25427} + 2.72\underline{07709}\ i$$

$$Z_2 = 3.14\underline{06429} + 2.71\underline{57922}\ i$$

$$Z_3 = 0.31\underline{530804} + 0.27\underline{323534}\ i$$

$$Z_4 = 0.31\underline{301054} + 0.27\underline{042248}\ i$$

（＿＿の部分は誤差）

絶対値の小さい解 Z_3, Z_4 は解 Z_1, Z_2 に比較して精度が悪い。

これに反し、次数の高い項の係数に誤差が多く入っていると

$$C_4 = 3.141\underline{4000} + 2.718\underline{1000}\ i$$

$$C_3 = -5.4571\underline{800} - 37.5748\underline{10}\ i$$

$$C_2 = -54.47384\underline{2} + 85.16337\underline{4}\ i$$

$$C_1 = 62.82212\underline{2} - 18.64123\underline{2}\ i$$

$$C_0 = -11.27425\underline{6} - 5.10022\underline{96}\ i$$

$$Z_1 = 3.1\underline{686124} + 2.7\underline{432769}\ i$$

$$Z_2 = 3.1\underline{149880} + 2.6\underline{936860}\ i$$

$$Z_3 = 0.314\underline{36722} + 0.272\underline{14006}\ i$$

$$Z_4 = 0.313\underline{95197} + 0.271\underline{51739}\ i$$

絶対値の大きい解 Z_1, Z_2 は解 Z_3, Z_4 に比較して精度が悪い。

ここで実際に（3−2）式を用いて高次の項より除算する.

$$\overline{C}_{n-1} = C_n$$

$$\overline{C}_{n-2} = C_n Z_1 + C_{n-1}$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$\overline{C}_1 = C_n Z_1^{n-2} + C_{n-1} Z_1^{n-3} + \cdots\cdots\cdots + C_3 Z_1 + C_2$$

$$\overline{C}_0 = C_n Z_1^{n-1} + C_{n-1} Z_1^{n-2} + \cdots\cdots\cdots + C_3 Z_1^2 + C_2 Z_1 + C_1$$

$$\overline{C}_{-1} = C_n Z_1^{n} + C_{n-1} Z_1^{n-1} + \cdots\cdots\cdots + C_3 Z_1^3 + C_2 Z_1^2 + C_1 Z_1 + C_0 = 0.0$$

$$(3-8)$$

（3−8）式の$\overline{C}_{-1}$は完全な零ではなく、収束判定（1−2）式より

$$|\pm\Delta \overline{C}_{-1}| = |\pm\Delta C_{i1} Z_1^{i1}| = max_{i=0,1,2,\cdots\cdots,n}(|\pm\Delta C_i Z_1^i|) \qquad (3-8')$$

程度の誤差をもち、又、（3−8）式の係数の誤差は

$$|\pm\Delta \overline{C}_i| = max_{j=i+1,\, i+2,\cdots\cdots,n}(|\pm\Delta C_j Z_1^{j-(i+1)}|)$$

$$i = 0,1,2,\cdots\cdots,n-1 \qquad (3-8'')$$

である。次に、（3−8）式の係数を用いて、（3−2）式によりZ_2を求めるのであるが、（3−1）式により求められるZ_2と比較して誤差はどうであろうか。誤差を比較するために、（3−1）式にZ_2を代入した時に得られる項と同じ項を含む次式を考えてみる。

$$\overline{f}(Z)\cdot Z_2 + \overline{C}_{-1} = 0 \qquad (3-9)$$

（3−9）式にZ_2を代入し、各項を（3−8）式で書き換え、

$$\overline{C}_{n-1} Z_2^n = C_n \cdot Z_2^n$$

$$\overline{C}_{n-2} Z_2^{n-1} = C_n Z_1 Z_2^{n-1} + C_{n-1} \cdot Z_2^{n-1}$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$\overline{C}_1 Z_2^2 = C_n Z_1^{n-2} Z_2^2 + C_{n-1} Z_1^{n-3} Z_2^2 + \cdots\cdots + C_2 \cdot Z_2^2$$

$$\bar{C}_0 Z_2 = C_n Z_1^{n-1} Z_2 + C_{n-1} Z_1^{n-2} Z_2 + \cdots\cdots + C_2 Z_1 Z_2 + C_1 \cdot Z_2$$
$$\bar{C}_{-1} = C_n Z_1^{n} + C_{n-1} Z_1^{n-1} + \cdots\cdots + C_2 Z_1^2 + C_1 Z_1 + C_0$$
$$(3-10)$$

(3−10)式の右辺の最後の項のみを集めると(3−1)式にZ_2を代入した場合の各項と同じになる．又(3−9)式では定数としてのZ_2を(3−2)式に乗じ，零(誤差)である$\bar{C}_{-1}$を加えている．従つて，もし(3−1)式でZ_1を求める時，無限桁の収束計算をすれば$\bar{C}_{-1}=0$であり，(3−1)式で求められるZ_2と，(3−9)式即ち(3−2)式で求められるZ_2は等しくなる。ところが，(3−2)式でZ_2を求める時の収束判定は

$$|\pm\Delta\bar{C}_{j2} Z_2^{j2+1}| = max_{j=0,1,2,\cdots,n}(|\pm\Delta\bar{C}_{j-1} Z_2^{j}|) \qquad (3-11)$$

であるから，$\bar{C}_{-1}$は零でなく誤差である。一方(3−1)式でZ_2を求める時の収束判定は

$$|\pm\Delta C_{i2} Z_2^{i2}| = max_{i=0,1,2,\cdots,n}(|\pm\Delta C_i Z_2^{i}|) \qquad (3-12)$$

である．ここで，(3−11)式と(3−12)式との間に

$$|\pm\Delta\bar{C}_{j2} Z_2^{j2+1}| > |\pm\Delta C_{i2} Z_2^{i2}| \qquad (3-13)$$

の関係があると，(3−2)式で得られるZ_2は，(3−1)式から得られるZ_2よりも大きい誤差を含む事になる。又，(3−8)式で定めた$\pm\Delta C_{i1}$を含む項をもつ(3−8)式で，誤差の絶対値が最大の項は$\pm\Delta C_{i1}$を含む項である。依つて(3−8′)式，(3−8″)式中$\pm\Delta C_{i1}$を含む項をもつ$|\pm\Delta\bar{C}_i|$は

$$|\pm\Delta\bar{C}_i| = max_{j=i+1,\,i+2,\cdots\cdots,\,n}\left(|\pm\Delta C_j Z_1^{j-(i+1)}|\right) = |\pm\Delta C_{i1} Z_1^{i1-(i+1)}|$$

$$i = -1,\ 0,\ 1,\cdots\cdots,\ i1-1 \qquad (3-14)$$

になる。Z_1とZ_2の関係を

$$Z_1 > Z_2 \qquad (3-15)$$

とすると、（3－7）式より $i1 \geq i2$ であるから $\pm\Delta C_{i2}$ を含む項をもつ（3－8）式は必ず $\pm\Delta C_{i1}$ を含む項をもつ。従って、$i1 > i2$ の場合は（3－14）式で $\pm\Delta C_{i2}$ を含む項は選ばれず、この $|\pm\Delta\bar{C}_i|$ を用いている（3－11）式では $\pm\Delta C_{i2}$ を含む項は勿論選ばれなく、（3－13）式の関係が成り立つ。又、$i1=i2$ の場合は（3－14）式を（3－11）式に代入し、$\pm\Delta C_{i1}=\pm\Delta C_{i2}$ を含む項の中で誤差の絶対値が最も大きい項を求めると

$$|\pm\Delta C_{i1} Z_1^{i1}| = max_{i=-1,0,1,\cdots\cdots,\,i1-1}\left(|\pm\Delta C_{i1} Z_1^{i1-(i+1)} Z_2^{i+1}|\right)$$

$$(3-16)$$

であるから、やはり（3－13）式の関係が成り立つ。即ち、

1. 得られた解の絶対値より絶対値の小さい解を、高次の項から（3－2）式を除して得られた $\bar{f}(Z)$ から求めると、直接 $f(Z)$ から得られる解よりも、解の精度が悪くなる。

次に、Z_1とZ_2の関係が

$$Z_1 \leq Z_2 \qquad (3-17)$$

である時はどうなるであろうか。（3－10）式中で誤差の絶対

値が最大の項は

$$\left|\pm\Delta C_i Z_2^i\right| = \max_{j=0,1,2,\cdots,i}\left(\left|\pm\Delta C_i Z_1^{i-j} Z_2^j\right|\right)$$

$$i = 0, 1, 2, \cdots\cdots, n \qquad (3-18)$$

の中にあるから(3−13)式は成立せず.(3−2)式で求められるZ_2が(3−1)式から求められるZ_2と同じ精度で得られる.次に例をあげる.(計算は10進法 8桁で行なう.)

$$f(Z) = \sum_{i=0}^{3} C_i Z^i = 0 \qquad (3-19)$$

$C_3 = 1.0000000$ $\qquad$ $C_1 = 9.9692877 \times 10^6$

$C_2 = -3.1733228 \times 10^4$ $\qquad$ $C_0 = -3.1006278 \times 10^7$

真の解は $Z_2 = \pi$, $Z_1 = \pi \times 10^2$, $Z_3 = \pi \times 10^4$

まず $Z_1 = 3.1415927 \times 10^2$ を求め.(3−8)式により

$\overline{C}_2 = C_3 \qquad\qquad = 1.0000000$

$\overline{C}_1 = C_3 Z_1 + C_2 \qquad = -3.1419069 \times 10^4$

$\overline{C}_0 = (C_3 Z_1 + C_2) Z_1 + C_1 = 9.8695900 \times 10^4$

次いで得られる解は

$Z_2 = 3.1415880$, $\qquad Z_3 = 3.1415927 \times 10^4$

Z_2の絶対値はZ_1の絶対値より小さいから解の精度は悪くなり、Z_3の絶対値はZ_1の絶対値より大きいから解の精度は悪くならない。

先に.(3−2)式を高次の項より除算したが.低次の項より除算すると

$$\bar{C}_n = (C_0/Z_1^n + C_1/Z_1^{n-1} + \cdots\cdots + C_{n-2}/Z_1^2 + C_{n-1}/Z_1 + C_n) = 0.0$$

$$\bar{C}_{n-1} = -(C_0/Z_1^n + C_1/Z_1^{n-1} + \cdots\cdots + C_{n-2}/Z_1^2 + C_{n-1}/Z_1)$$

$$\bar{C}_{n-2} = -(C_0/Z_1^{n-1} + C_1/Z_1^{n-2} + \cdots\cdots + C_{n-2}/Z_1)$$

$$\vdots$$

$$\bar{C}_1 = -(C_0/Z_1^2 + C_1/Z_1)$$

$$\bar{C}_0 = -(C_0/Z_1) \tag{3-20}$$

(3−20)式の $\bar{C}_n$ は完全な零ではなく、(1−2)式より

$$|\pm\Delta\bar{C}_n| = |\pm\Delta C_{i1} Z_1^{i1-n}| = max_{i=0,1,2,\cdots,n}(|\pm\Delta C_i Z_1^{i-n}|) \tag{3-20'}$$

程度の誤差をもち、(3−20)式の係数の誤差は

$$|\pm\Delta\bar{C}_i| = max_{j=0,1,2,\cdots,i}(|\pm\Delta C_j Z_1^{j-(i+1)}|)$$

$$i = 0, 1, 2, \cdots\cdots, n-1 \tag{3-20''}$$

である。次に、(3−9)式と同じ理由で

$$f(Z)\cdot Z_1 + \bar{C}_n Z_2^n = 0 \tag{3-21}$$

(3−21)式を考え、(3−20)式の係数を用いて書き換える。

$$\bar{C}_n Z_2^n = (C_0 Z_2^n/Z_1^n + C_1 Z_2^n/Z_1^{n-1} + \cdots\cdots + C_{n-2} Z_2^n/Z_1^2 + C_{n-1} Z_2^n/Z_1 + C_n Z_2^n)$$

$$\bar{C}_{n-1} Z_2^{n-1} Z_1 = -(C_0 Z_2^{n-1}/Z_1^{n-1} + C_1 Z_2^{n-1}/Z_1^{n-2} + \cdots\cdots + C_{n-2} Z_2^{n-1}/Z_1 + C_{n-1} Z_2^{n-1})$$

$$\bar{C}_{n-2} Z_2^{n-2} Z_1 = -(C_0 Z_2^{n-2}/Z_1^{n-2} + C_1 Z_2^{n-2}/Z_1^{n-3} + \cdots\cdots + C_{n-2} Z_2^{n-2})$$

$$\overline{C}_1 \; Z_2 \; Z_1 = -(C_0 Z_2 / Z_1 + C_1 Z_2)$$
$$\overline{C}_0 \qquad Z_1 = -(C_0)$$

(3−22)

(3−22)式の右辺の最後の項のみを集めると(3−1)式にZ_2を代入した場合の各項と同じになる。(3−22)式の係数を用いた場合、(3−2)式でZ_2を求める時の収束判定は

$$|\varepsilon| = max\{ max_{j=0,1,2,\cdots,n-1}(|\pm\Delta\overline{C}_j Z_2^j Z_1|), |\pm\Delta\overline{C}_n Z_2^n| \}$$

(3−23)

である。(3−23)式と(3−12)式との間に

$$|\varepsilon| > |\pm\Delta C_{i2} Z_2^{i2}| \qquad (3-24)$$

の関係があれば、(3−2)式で得られるZ_2は、(3−1)式から得られるZ_2よりも大きい誤差を含む。一方、(3−20′)式で定めた$\pm\Delta C_{i1}$を含む項をもつ(3−20)式で、誤差の絶対値が最大の項は$\pm\Delta C_{i1}$を含む項である。依って(3−20′)式の中で$\pm\Delta C_{i1}$を含む項をもつ$|\pm\Delta\overline{C}_i|$は

$$|\pm\Delta\overline{C}_i| = max_{j=0,1,2,\cdots,i}(|\pm\Delta C_j Z_1^{j-(i+1)}|) = |\pm\Delta C_{i1} Z_1^{i1-(i+1)}|$$
$$i = i1,\ i1+1, \cdots, n-1 \qquad (3-25)$$

になる。Z_1とZ_2の関係を

$$Z_1 < Z_2 \qquad (3-26)$$

とすると、(3−7)式より$i1 \leq i2$であるから$\pm\Delta C_{i2}$を含む

項をもつ(3−20)式は必ず$\pm\Delta C_{i1}$を含む項をもつ。従って、$i1<i2$の場合は(3−25)式で$\pm\Delta C_{i2}$を含む項は選ばれず、この$|\pm\Delta \overline{C}_i|$を用いている(3−23)式では$\pm\Delta C_{i2}$を含む項は勿論選ばれなく、(3−24)式が成り立つ。又、$i1=i2$の場合は(3−25)式を(3−23)式に代入し、$\pm\Delta C_{i1}=\pm\Delta C_{i2}$を含む項の中で誤差の絶対値が最も大きい項を求めると

$$\left|\pm\Delta C_{i1} Z_2^n / Z_1^{n-i1}\right| = max_{i=i1,\, i1+1,\,\cdots\cdots,\, n}\left(\left|\pm\Delta C_{i1} Z_2^i / Z_1^{i-i1}\right|\right) \qquad (3-27)$$

であるから、やはり(3−24)式の関係が成り立つ。即ち

2. 得られた解の絶対値より絶対値の大きい解を、低次の項から(3−2)式を除して得られた$\tilde{f}(Z)$から求めると、直接$f(Z)$から得られる解よりも、解の精度が悪くなる。

次に、Z_1とZ_2の関係が

$$Z_1 \geq Z_2 \qquad (3-28)$$

である時はどうなるであろうか。(3−22)式中で誤差の絶対値が最大の項は

$$\left|\pm\Delta C_i Z_2^i\right| = max_{j=i,\, i+1,\,\cdots\cdots,\, n}\left(\left|\pm\Delta C_i Z_2^j / Z_1^{j-i}\right|\right)$$
$$i=0,1,2,\cdots\cdots,n \qquad (3-29)$$

の中にあるから(3−24)式は成立せず、(3−2)式で求められるZ_2が(3−1)式から求められるZ_2と同じ精度で得られる。次に、(3−19)式を用いた例をあげる。

$$\overline{C}_2 = -[(C_0/Z_1 + C_1)/Z_1 + C_2]/Z_1 = 9.9999914 \times 10^{1}$$
$$\overline{C}_1 = -(C_0/Z_1 + C_1)/Z_1 = -3.1419069 \times 10^{4}$$
$$\overline{C}_0 = -(C_0/Z_1) = 9.8696047 \times 10^{4}$$
$$Z_2 = 3.1415927, \qquad Z_3 = 3.1415954 \times 10^{4}$$

Z_2の絶対値はZ_1の絶対値より小さいから解の精度は悪くならない。Z_3の絶対値はZ_1の絶対値より大きいから解の精度は悪くなる。

しからば、(3−2)式の係数は、どうして計算したらZ_1の誤差を入れずに求められるか。これには2通りの計算方法が考えられる。その1つは非常に多くの桁をとり、(3−1)式でZ_1を求める時に収束の桁数を多くする方法であり、他の1つは高次の項、低次の項の両方から($Z-Z_1$)で除算し、(3−3)式により得られるi1の項をその境とする方法である。では、第1の方法：桁を多くとつて、収束桁数を多くする方法について述べる。まず、桁数をどれだけとればよいか。(3−1)式を用いてZ_1を求める時、高次の項より除算して(3−2)式の係数を求めるとすれば、Z_1の絶対値より絶対値の小さい解についてのみ考えればよい。今、(3−1)式でZ_2を求めると、収束判定は(3−12)式であるから、(3−10)式の各項の計算を行なう時、その各項に入る4捨5入による誤差が(3−12)式より小さくなれば、Z_2の誤差は、(3−2)式でZ_2

を求めても，（3－1）式でZ_2を求めた時より大きくはならない。又，Z_1を求める時，他の解は未知であるから（3－12）式の値はわからない。そこで，（3－12）式で$|\pm\Delta C_0|$が選ばれたとすると，（3－10）式のすべての項に入る4捨5入の誤差は$|\pm\Delta C_0|$より小さくなる様に桁をとらなければならない事になる。実際に，Z_1を求めた後でZ_2が求められ，

$$\left|\pm\Delta C_{i2} Z_2^{i2}\right| > \left|\pm\Delta C_0\right| \qquad (3-30)$$

であったとしても，$|Z_2| < |Z_1|$であるから

$$\left|\pm\Delta C_i Z_1^i\right| = \max_{j=0,1,2,\cdots\cdots,i}\left(\left|\pm\Delta C_i Z_1^j Z_2^{i-j}\right|\right)$$

$$i = 0, 1, 2, \cdots\cdots, n \qquad (3-31)$$

が成立し，$C_i Z_1^i$（$i=0,1,\cdots\cdots,n$）に入る4捨5入の誤差が$|\pm\Delta C_0|$より小さい様に，C_i，Z_1^i（$i=0,1,\cdots\cdots,n$）の桁をとっているから，当然$|\pm\Delta C_{i2} Z_2^{i2}|$より小さい4捨5入の誤差しか（3－10）式のすべての項には入らない。

次に，桁数を多くとつて計算した例を示す。

①係数の精度は4桁　③ ＿＿ の部分は誤差

②計算桁数は10桁　④ ΔΔΔ の部分は4捨5入による誤差

$\sum_{i=0}^{3} C_i Z^i = 0$　　真の解は$Z_1 = \pi\times10^4$，$Z_2 = \pi\times10^2$，$Z_3 = \pi$

$C_3 = 1.000\underline{000000}$　　　$C_1 = 9.969\underline{000000}\times10^6$

$C_2 = -3.173\underline{000000}\times10^4$　　　$C_0 = -3.100\underline{000000}\times10^7$

まず，Z_1を10桁計算で求める。$Z_1 = 3.141\underline{267545}\times10^4$

$$\overline{C}_2 = C_3$$

$$\overline{C}_1 = C_3 Z_1 + C_2 = \overline{C}_2 Z_1 + C_2$$

$$\overline{C}_0 = (C_3 Z_1 + C_2) Z_1 + C_1 = \overline{C}_1 Z_1 + C_1$$

$$\overline{C}_{-1} = [(C_3 Z_1 + C_2) Z_1 + C_1] Z_1 + C_0 = \overline{C}_0 Z_1 + C_0$$

$$\overline{C}_2 Z_1 = 3.141\underline{267545} \times 10^4$$

$$C_2 = -3.173\underline{000000} \times 10^4$$

$$\overline{C}_1 = 3.1\underline{732455\overset{\triangle}{0}\overset{\triangle}{0}} \times 10^2$$

$$\overline{C}_1 Z_1 = -9.9\underline{680131\overset{\triangle}{0}\overset{\triangle}{1}} \times 10^6$$

$$C_1 = 9.969\underline{000000} \times 10^6$$

$$\overline{C}_0 = \underline{9.868\overset{\triangle}{9}\overset{\triangle}{9}\overset{\triangle}{0}\overset{\triangle}{0}\overset{\triangle}{0}\overset{\triangle}{0}} \times 10^2$$

$$\overline{C}_0 Z_1 = \underline{3.100\overset{\triangle}{1}\overset{\triangle}{1}\overset{\triangle}{3}\overset{\triangle}{7}\overset{\triangle}{9}\overset{\triangle}{9}} \times 10^7$$

$$C_0 = -3.100\underline{000000} \times 10^7$$

$$\overline{C}_{-1} = \underline{\overset{\triangle}{1}.\overset{\triangle}{1}\overset{\triangle}{3}\overset{\triangle}{7}\overset{\triangle}{9}\overset{\triangle}{9}\overset{\triangle}{0}\overset{\triangle}{0}\overset{\triangle}{0}\overset{\triangle}{0}} \times 10^3$$

この様に$\overline{C}_0 Z_1$にC_0を加算する時、C_0の ＿ の桁に$\overline{C}_0 Z_1$の△△の桁が入るように桁数をとつて収束計算を行なえば、解$Z_2 = \pi \times 10^2$，$Z_3 = \pi$は$Z_1 = \pi \times 10^4$の解を除いた2次方程式で解いても、初めに与えられた3次代数方程式から直接解かれた解Z_2，Z_3と精度は変らない。

ベアストー法（ヒッチコック法）で収束計算を行なう時にも、同様に桁数を増加させないと、次数の下げられた代数方程式より得られる解は、初めに与えられた代数方程式から直

接求めた解よりも誤差は大きくなる．次に、4次代数方程式の例を示す。

①係数の精度は4桁　③ —— の部分は誤差

②計算桁数は9桁　④ △△△ の部分は4捨5入による誤差

$$\sum_{i=0}^{4} C_i Z^i = 0$$

真の解は $Z_1 = \pi \times 10^3$， $Z_2 = \pi \times 10^2$， $Z_3 = \pi \times 10$， $Z_4 = \pi$

$C_4 = 3.141\underline{00000}$　　　$C_2 = 3.476\underline{00000} \times 10^6$

$C_3 = -1.097\underline{00000} \times 10^4$　　　$C_1 = -1.082\underline{00000} \times 10^8$

$C_0 = 3.060\underline{00000} \times 10^8$

まず，Z_1 と Z_2 を9桁計算で求め，$Z^2 - (Z_1 + Z_2)Z + Z_1 Z_2$ で除す。$(Z_1 + Z_2) = 3.457\underline{96527} \times 10^3$， $Z_1 Z_2 = 9.870\underline{72031} \times 10^5$

$\bar{C}_2 = C_4$　　　$\bar{\bar{C}}_2 = C_3$

$\bar{C}_1 = \bar{C}_2 (Z_1 + Z_2) + \bar{\bar{C}}_2$　　　$\bar{\bar{C}}_1 = C_2 - \bar{C}_2 Z_1 Z_2$

$\bar{C}_0 = \bar{C}_1 (Z_1 + Z_2) + \bar{\bar{C}}_1$　　　$\bar{\bar{C}}_0 = C_1 - \bar{C}_1 Z_1 Z_2$

$\bar{C}_{-1} = \bar{C}_0 (Z_1 + Z_2) + \bar{\bar{C}}_0 = 0.0$　　　$\bar{\bar{C}}_{-1} = C_0 - \bar{C}_0 Z_1 Z_2 = 0.0$

$$\begin{array}{rl} \bar{C}_2 (Z_1 + Z_2) = & 1.086\underline{14689} \times 10^4 \\ \bar{\bar{C}}_2 = & -1.097\underline{00000} \times 10^4 \\ \hline \bar{C}_1 = & 1.0\underline{85311\overset{\triangle}{0}\overset{\triangle}{0}} \times 10^2 \end{array}$$

$$\begin{array}{rl} C_2 = & 3.476\underline{00000} \times 10^6 \\ -\bar{C}_2 Z_1 Z_2 = & -3.100\underline{39325} \times 10^6 \\ \hline \bar{\bar{C}}_1 = & 3.75\underline{60675\overset{\triangle}{0}} \times 10^5 \end{array}$$

$$\bar{C}_1(Z_1+Z_2) = -3.7\underline{5296775}\times 10^5$$
$$\bar{\bar{C}}_1 = 3.75\underline{606750}\times 10^5$$
$$\bar{C}_0 = \underline{3.09975000}\times 10^2$$

$$C_1 = -1.082\underline{00000}\times 10^8$$
$$-\bar{C}_1 Z_1 Z_2 = 1.07\underline{128013}\times 10^8$$
$$\bar{\bar{C}}_0 = -\underline{1.07198700}\times 10^6$$

$$\bar{C}_0(Z_1+Z_2) = \underline{1.07188278}\times 10^6$$
$$\bar{\bar{C}}_0 = -\underline{1.07198700}\times 10^6$$
$$\bar{C}_{-1} = -\underline{1.04220000}\times 10^2$$

$$C_0 = 3.06\underline{000000}\times 10^8$$
$$-\bar{C}_0 Z_1 Z_2 = -\underline{3.05967653}\times 10^8$$
$$\bar{\bar{C}}_{-1} = \underline{3.23470000}\times 10^4$$

C_0に$-\bar{C}_0 Z_1 Z_2$を加算する時、C_0の—の桁に$-\bar{C}_0 Z_1 Z_2$の△△の桁が入る様に桁数をとつて収束計算をすれば、次数の下げられた2次方程式から得られる解Z_3, Z_4は、初めに与えられた4次方程式から直接求めた解Z_3, Z_4と同じ精度が得られる。

最後に、桁数を増加させずに、且つ得られた解の誤差を入れずに、次数を1つ下げた代数方程式の係数の求め方を述べる。これが、数値解法が有限桁である事をはつきり示している点であり、単なる多項式の除算であつても計算する順序が異なると、同じ式で計算しても解が得られない事がある。

得られた解Z_1を用いて

$$|\pm\Delta C_{imax} Z_1^{imax}| = \max_{i=0,1,2,\cdots,n} (|\pm\Delta C_i Z_1^i|)$$

(3−32)

なる$imax$を求め このimaxを用いて（3−1）式を

$$f(Z) = \left(\sum_{i=0}^{imax-1} C_i Z^i\right) + C_{imax} Z^{imax} + \left(\sum_{j=imax+1}^{n} C_j Z^j\right) \qquad (3-33)$$

の様に分離し（3−9）式の第1項は$(Z-Z_1)$で低次の項より除し 第3項は$(Z-Z_1)$で高次の項より除す.（3−2）式中の$\bar{C}_i$は

$$\begin{aligned}
\bar{C}_{n-1} &= C_n \\
\bar{C}_{n-2} &= C_n Z_1 + C_{n-1} \\
&\vdots \\
\bar{C}_{imax} &= C_n Z_1^{n-(imax+1)} + C_{n-1} Z_1^{n-(imax+2)} + \cdots\cdots + C_{imax+1} \\
\bar{C}_{imax-1} &= -(C_0/Z_1^{imax} + C_1/Z_1^{imax-1} + \cdots\cdots + C_{imax-1}/Z_1) \\
&\vdots \\
\bar{C}_1 &= -(C_0/Z_1^2 + C_1/Z_1) \\
\bar{C}_0 &= -\ C_0/Z_1 \\
\bar{C}_{-1} &= C_n Z_1^{n-imax} + C_{n-1} Z_1^{n-(imax+1)} + \cdots\cdots + C_{imax+1} Z_1 \\
&\quad + C_{imax} \\
&\quad + C_0/Z_1^{imax} + C_1/Z_1^{imax-1} + \cdots\cdots + C_{imax-1}/Z_1 \\
&= f(Z_1)/Z^{imax} = 0.0
\end{aligned}$$

(3−34)

$\bar{C}_{-1}$ は低次の項、高次の項の両側より除して、中央に残る項である。ところで（3−2）式で次の解Z_2を求めようとすると（3−2）式に定数を乗じ、零である項を加えた式でZ_2を求めても、（3−2）式から直接得られる解Z_2と等しい解が得られる筈であるから、（3−2）式の代りに次の式で解を求める。

$$|Z_2| \geq |Z_1| \qquad \bar{f}(Z) \times Z_2 + \bar{C}_{-1} Z_2^{imax} = 0 \qquad (3-35)$$

$$|Z_2| < |Z_1| \qquad \bar{f}(Z) \times Z_1 + \bar{C}_{-1} Z_2^{imax} = 0 \qquad (3-36)$$

（3−35）式の第1項でのZ_2及び（3−36）式の第1項でのZ_1は定数と考え、$\bar{C}_{-1} = 0$であるから（3−35）、（3−36）式の第2項は零である。ここでは（3−36）式の説明のみ行なうが、（3−35）式についても同様に説明が出来る。まず（3−36）式をZ_1、Z_2及び（3−1）式の係数C_iを用いて書き直してみる。

$$\begin{aligned}
\bar{C}_{n-1} Z_2^{n-1} Z_1 &= C_n Z_2^{n-1} Z_1 \\
\bar{C}_{n-2} Z_2^{n-2} Z_1 &= C_n Z_2^{n-2} Z_1^2 + C_{n-1} Z_2^{n-2} Z_1 \\
&\vdots \\
\bar{C}_{imax} Z_2^{imax} Z_1 &= C_n Z_2^{imax} Z_1^{n-imax} + C_{n-1} Z_2^{imax} Z_1^{n-(imax+1)} + \cdots \\
&\quad \cdots + C_{imax+1} Z_2^{imax} Z_1 \\
\bar{C}_{-1} Z_2^{imax} &= C_n Z_2^{imax} Z_1^{n-imax} + C_{n-1} Z_2^{imax} Z_1^{n-(imax+1)} + \\
&\quad \cdots + C_{imax+1} Z_2^{imax} Z_1 + C_{imax} Z_2^{imax} \\
&\quad + C_0 Z_2^{imax} / Z_1^{imax} + C_1 Z_2^{imax} / Z_1^{imax-1} + \cdots \\
&\quad \cdots + C_{imax-1} Z_2^{imax} / Z_1
\end{aligned}$$

$$\bar{C}_{imax-1} Z_2^{imax-1} Z_1 = -(C_0 Z_2^{imax-1} / Z_1^{imax-1} + C_1 Z_2^{imax-1} / Z_1^{imax-2} + \cdots\cdots + C_{imax-1} Z_2^{imax-1})$$

$$\vdots$$

$$\bar{C}_1 \quad Z_2 \quad Z_1 = -(C_0 Z_2 / Z_1 \quad + C_1 Z_2)$$

$$\bar{C}_0 \qquad Z_1 = - \ C_0 \qquad (3-37)$$

(3−37)の各式の右辺の合計を零とする様にZ_2を求めると、(3−36)式即ち(3−2)式で解Z_2を求める事と同じになる。従って(3−37)の各式の右辺で誤差の最も大きい項を求め、その誤差が(3−1)式で解Z_2を求める時の最も誤差の大きい項の誤差と等しい数値であるならば、(3−1)式で求める解Z_2と、(3−2)式で求める解Z_2の精度が同じわけである。(3−13)式の中で$\pm\Delta C_i$ ($i = n, n-1, \cdots\cdots imax$)による最大の誤差はどうであろうか。それは$|Z_2| < |Z_1|$であり、且つ(3−32)式より$|\pm\Delta C_{imax} Z_2^{imax}|$であることが知れる。又$\pm\Delta C_i$ ($i = 0, 1, 2, \cdots\cdots, imax-1$)による最大の誤差は

$$max_{i=0,1,2,\cdots\cdots,imax-1} (|\pm\Delta C_i Z_2^i|)$$

であるから、(3−37)の各式の中で最も誤差の大きい項は

$$max_{i=0,1,2,\cdots\cdots,imax} (|\pm\Delta C_i Z_2^i|)$$

一方、(3−1)式にZ_2を代入して、最も誤差の大きい項は

$$max_{i=0,1,2,\cdots\cdots,n} (|\pm\Delta C_i Z_2^i|)$$

であり、$|Z_2| < |Z_1|$であるから(3−7)式より両者は等しい

。次に（3　19）式を用いた例をあげる.

$\bar{C}_2 = C_3 \qquad = 1.0000000$

$\bar{C}_1 = C_3 Z_1 + C_2 = 3.1419069 \times 10^4$

$\bar{C}_0 = -C_0 / Z_1 \quad = 9.8696047 \times 10^4$

$Z_2 = 3.1415927$,　　$Z_3 = 3.1415927 \times 10^4$

Z_2, Z_3共に解の精度は悪くならない。